ACRM GROUP ASSIGNMENT

DESERT SELLING COMPANY

2021-2022

1 Practicalities

1.1 Form groups

- Form groups of **4 people**. Subscribe as a team to a group on Ufora (Groups > Group Assignment) **before the 3th of October at 12:00 am**. Students who have not enrolled by then will be automatically and randomly assigned to a group.

1.2 Deadline

- The deadline of the group assignment is the **24th of October (12 am)**. By then, you should upload your PowerPoint presentation on Ufora (Ufora tools > Assignments > Group Assignment) and are not allowed to make any changes to your code on GitHub (see later) anymore.
- The presentations will be held online and will take place between the **25th and 29th of October**. The exact schedule will follow later.

1.3 Presentation

- At the end of the assignment each group will present their findings within 20 min.
  - **10 min presentation**
  - **8 min Q&A**
  - **2 min Feedback**

- Create a **professional** presentation
  - Presentation should be directed at a non-technical audience. However, questions might be technical as well.
  - Present as if you would present to the management of the desert selling company. Try to interpret your findings and communicate a message that is relevant to them.
  - Make it visually appealing

- There will be strong time pressure:
  - Knowledge to do the analyses is taught during classes
  - Presentation time is limited, so it is necessary to select the **most relevant results**, without losing information

    - E.g. Often different plots can be combined into a single plot, saving space and presentation time and making comparison easier
    - Having backup slides with additional information might be a good idea…

- Do not limit yourselves to the obvious outcomes. If you know additional analyses that might be relevant and that are not mentioned in the goals below, feel free to do it. Moreover, go beyond simply reporting results, interpret them as well and search for links.

- Try to come with a **clear message** during your presentation.

- You will have to create your solution using a **GitHub** repository. GitHub is an online platform to facilitate coding projects which is very popular in the IT/Data Science industry. You will have an on campus introductory course on how to use GitHub. In the meantime, you are of course allowed to work outside GitHub. Once you create your repository on github.ugent.be, you should add Professor Van den Poel (dvdpoel), Bram (banjanss), Lisa (lschetge), and Arno (aliseune) to the repository. The final solution should be included in the repository by the end of the project.

2 Assignment

For this project, data was collected from a **company that sells a various number of desserts at home**. The data was collected from January 2019 to December 2021 and contains different data tables containing information regarding the customers, employees, transactions, routes, etc. The **goal of the assignment is to analyze how the company can improve its services**. In particular, the company wants to gain more insights in which customers add the most value to the company, which customers left the company and can be won back, which employees render the highest turnover, which routes should be adjusted, and which routes should be reassigned to different depots. The management of the company expects **a clear presentation in which the different analyses are discussed, and which actions the company should take to improve its services**.

Q1. What are the products?

- What are the most frequently bought products?
- Which products render the most revenue?
- Which products are bought the most in the region of Brussels, Antwerp, …?
- Are product purchases correlated? Are some products often purchased together?
- What are the total sales generated for each product family?
- Does the weather / seasonal changes have an effect on the total revenue?

Q2. Who are the customers?

- What do the customers buy?
- Which customers left the company?
- Which customers have the highest CLV?

- What is the relationship between leaving the company and buying patterns?
- Are there clusters of customers? How would you describe these different clusters?
- Do customers have different buying patterns during the weekend?
- Do customer sales differ across different cities? Is there a relationship between customer sales and average income per inhabitant (and other factors)?

Q3. Who are the employees?

- What are the routes of the employees?
- What is the turnover for each employee?

Q4. What are the routes?

- How are the customers divided into regions?
- Which customers are assigned to which routes?
- Which routes are assigned to which depots?

Q5. How can the company improve its service?

- Which customers should be rewarded?
- Which employees should be rewarded?
- To which customers should the company send coupons in order to win them back?
- Are there factors that the company can change in order to decrease the churning rate?
- Would it be valuable to recommend (upsell / cross sell) products to a customer?
- Which employees should be assigned to different routes?
- Which routes should be reassigned to different depots?
- Which customers should be reassigned to different routes?
- Which depots should be removed? Where should the company create new depots?
- Which products should be added / removed from depots?
- Does customer satisfaction relate to different factors? Can the company respond to these factors?

Overall, your presentation should make a comprehensive summary of the requested items. If you feel that one of the points requested above does not fit in your presentation, you can leave them out (or add it as extra slides). However, **you need to write code for most of the questions above**. There are a lot of extra aspects that can be looked at. You can include these in your presentation or coding if you think they may be of particular relevance or contain some unexpected results.

NOTE: We will also ask you to hand in your Python code, please structure it in a readable format using comments and numbering as below for each exercise. Any additional code you write, should be indicated as “extra”

3 Hints and guidelines

- New functions that have not been seen in class might be necessary, so use the Internet wisely.
  - Google
  - Stack Overflow
  - Quora
  - Etc.
- Handle missing values (delete them or find a way to integrate them in your analyses).
- **Do not limit yourselves to the questions asked above.** This will only result in very moderate grades. Try to think further and compile a value-adding business case.
- Search for relevant datasets on the internet if necessary. (Infrabel, NMBS, ...)
- Use plots and make your slides visually appealing.
- Other information that you can find on the Internet can be useful to provide more detailed information.